

PERBANDINGAN ARSITEKTUR CONVOLUTIONAL NEURAL NETWORK UNTUK KLASIFIKASI FUNDUS

Wahyudi Setiawan¹⁾

¹ Jurusan Teknik Informatika, Fakultas Teknik, Universitas Trunojoyo
Jl. Raya Telang, PO BOX 2 Kamal, Bangkalan
E-mail : ¹wsetiawan@trunojoyo.ac.id

ABSTRAK

Pada artikel ini membahas tentang perbandingan arsitektur Convolutional Neural Network (CNN) untuk klasifikasi citra fundus. Arsitektur CNN yang diujicobakan yaitu AlexNet, Visual Geometry Group (VGG) 16, VGG19, Residual Network (ResNet) 50, ResNet101, GoogleNet, Inception-V3, InceptionResNetV2 dan Squeezenet. Citra ujicoba menggunakan fundus retina untuk mengklasifikasi 2 kelas yaitu normal dan neovaskularisasi. Citra dilakukan preprosesing yaitu dengan membaginya menjadi 16 bagian yang sama. Skenario ujicoba menggunakan 2 tahap yaitu, pertama, menggunakan CNN tanpa optimasi tambahan, kedua, CNN menggunakan optimasi Gradient Descent. Hasil ujicoba pada kedua skenario menunjukkan arsitektur terbaik yaitu VGG19 dan VGG16. Ujicoba tahap pertama menghasilkan sensitivitas, spesifisitas dan akurasi yaitu 87,8%, 90,7% dan 89,3%. Untuk ujicoba tahap kedua sensitivitas, spesifisitas dan akurasi yaitu 94,2%, 90,4% dan 92,31%.

Kata kunci : Klasifikasi, Citra Fundus, *Convolutional Neural Network*, *Visual Geometry Group (VGG)*, *Gradient Descent*.

ABSTRACT

This article discusses about comparison of Convolutional Neural Network (CNN) architecture for fundus image classification. The CNN architecture are AlexNet, Visual Geometry Group (VGG) 16, VGG19, Residual Network (ResNet) 50, ResNet101, GoogleNet, Inception-V3, InceptionResNetV2 and Squeezenet. The dataset of retinal fundus classify 2 classes: normal and neovascularization. The image is preprocessed by dividing it into 16 equal parts. The test scenario uses 2 steps, first, using CNN without additional optimization, second, CNN uses Gradient Descent optimization. The results of scenarios show the best architecture, VGG19 and VGG16. The first phase produced sensitivity, specificity and accuracy of 87.8%, 90.7% and 89.3%. For the second phase had sensitivity, specificity and accuracy of 94.2%, 90.4% and 92.31%.

Keywords: *Classification, Fundus Image, Convolutional Neural Network, Visual Geometry Group (VGG), Gradient Descent.*

PENDAHULUAN

CNN merupakan metode *supervised* deep learning. Saat pertama kali ada di publik dengan arsitektur LeNet tahun 1998, CNN tidak terlalu mendapatkan perhatian atau kurang sukses disebabkan keterbatasan yang ada diantaranya dataset ujicoba kecil, komputer yang lambat dan salah dalam menerapkan fungsi non linear [1].

Selanjutnya pada tahun 2012, terdapat arsitektur CNN AlexNet yang berhasil menjadi pemenang ImageNet *Competition* yaitu kompetisi untuk klasifikasi dan deteksi citra yang terdiri dari jutaan citra dengan puluhan ribu kelas. AlexNet terdiri dari total 25 layer dengan 8 *convolutional layer* dan 60 juta parameter [2]. Sukses yang telah dicetak oleh AlexNet diikuti dengan munculnya arsitektur atau model lain dari CNN diantaranya VGG [3], GoogleNet [4], ResNet [5], Inception-V3 [4], InceptionResNet-V2 [6] dan SqueezeNet [7].

CNN memiliki fungsi untuk melakukan ekstraksi fitur. Fitur-fitur perlu didapatkan guna proses atau tugas seperti klasifikasi, *clustering* ataupun regresi. Pada *machine learning* konvensional dilakukan ekstraksi fitur manual, artinya ditentukan terlebih dahulu fitur-fitur yang diekstraksi. Sedangkan CNN melakukan ekstraksi fitur secara otomatis pada *convolutional layer*, *pooling layer* dan juga aktivasi *Rectified Linear Unit* (ReLU). Selanjutnya fitur-fitur dilakukan proses klasifikasi pada *Fully Connected layer* (FCL) dan aktivasi softmax [8].

Artikel ini membahas tentang klasifikasi citra fundus untuk mengidentifikasi neovaskularisasi atau pembuluh darah baru pada retina menggunakan CNN. Sebelumnya telah dilakukan penelitian tentang klasifikasi citra fundus menggunakan CNN diantaranya yaitu Yu *et al.* melakukan klasifikasi citra fundus 2 kelas menggunakan arsitektur CNN 16 layer termasuk input dan output layer. Layer yang berada arsitektur CNN yang

digunakan diantaranya 7 Convolutional Layer, 4 max-pooling layer, 2 FCL, dan *dropout* 0.5. Hasil ujicoba mendapatkan akurasi 91,92%, Sensitivitas 88,85% dan Spesifisitas 96% [9]. Penelitian selanjutnya dilakukan oleh Kwasigroch *et al.* menggunakan CNN dengan arsitektur VGG-D yang terdiri dari 16 layer yaitu 5 Convolutional Layer, 5 max-pooling layer, 2 Drop-Out, *Dense Layer* dan *Softmax*. Hasil ujicoba menunjukkan akurasi 82% [10]. Penelitian selanjutnya yaitu Ghosh *et al.*, menggunakan CNN 29 layer yang terdiri dari 13 Convolutional layer, 5 max-pooling, 2 dropout, dense layer dan softmax [11].

Ragam arsitektur CNN dan dataset yang digunakan untuk proses ujicoba menyulitkan untuk dilakukan perbandingan. Pada penelitian ini dilakukan klasifikasi citra fundus untuk 2 kelas yaitu normal dan neovaskularisasi. Dataset yang diujicobakan yaitu *Retina Image Bank* dan MESSIDOR [12]. Kebaruan dari penelitian ini yaitu pertama melakukan perbandingan klasifikasi menggunakan 9 arsitektur CNN. Kedua menerapkan optimasi menggunakan algoritma gradient descent. Hasil penelitian menunjukkan arsitektur CNN yang paling direkomendasikan untuk melakukan klasifikasi sesuai dengan dataset ujicoba.

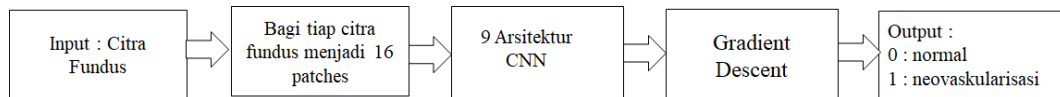
METODE

Tahap pertama, klasifikasi citra fundus menggunakan 9 arsitektur CNN yaitu AlexNet, VGG16, VGG19, ResNet50, ResNet101, GoogLeNet, Inception-V3, Inception-ResNetV2, dan Squeezenet. Fitur-fitur yang dihasilkan pada arsitektur CNN tersebut, selanjutnya dilakukan klasifikasi.

Tahap kedua, klasifikasi citra fundus menggunakan metode optimasi. Pada tahap ini dilakukan klasifikasi citra fundus menggunakan 9 arsitektur CNN yang sama dengan tahap pertama. Pada tahap ini juga diterapkan metode optimasi untuk meningkatkan prosentase akurasi. Metode optimasi yang digunakan

adalah *gradient descent* (GD). Metode ini merupakan metode optimasi paling sederhana secara komputasi dan memiliki hasil optimal pada saat diterapkan pada CNN. Optimasi GD dapat dilakukan dengan dua cara melalui momentum dan *adaptive subgradient*. Optimasi melalui momentum dilakukan dengan metode *Gradient Descent with Momentum*

(GDM) [13]. Optimasi melalui *adaptive subgradient* dilakukan dengan metode *Root Mean Square Propagation* (RMSProp) [14]. Optimasi melalui kombinasi keduanya dilakukan dengan metode *Adaptive Moment Optimization* (Adam) [15]. Tahapan ujicoba selengkapnya ditunjukkan pada Gambar 1.



Gambar 1. Tahapan klasifikasi citra fundus

Fundus yang digunakan untuk ujicoba terdiri dari 128 citra, selanjutnya dibagi menjadi 16 bagian yang sama sehingga menghasilkan 2048 patches.

HASIL DAN PEMBAHASAN

Pengukuran performa selengkapnya untuk arsitektur CNN terdapat pada Tabel 1. Hasil ujicoba menunjukkan akurasi tertinggi pada arsitektur VGG19 adalah 89,3%. Arsitektur VGG19 merupakan arsitektur yang terdiri dari 47 layer yang menggunakan konsep semakin banyak layer akurasi semakin baik.

Convolutional layer menggunakan filter 3×3 dengan jumlah dimensi bertambah di setiap layer. Dimensi yang digunakan pada filter adalah 3, 64, 128, 256, dan 512.

Tabel 1. Pengukuran performa klasifikasi citra retina menggunakan arsitektur CNN

Arsitektur CNN	Sens. (%)	Spec. (%)	Acc. (%)
AlexNet	88,8	87,1	87,8
VGG16	88,8	88,8	88,8
VGG19	90,4	88,2	89,3
GoogLeNet	87,7	85,4	86,5
Inception-V3	84,0	83,0	83,5
ResNet50	86,9	89,1	88,0
ResNet101	87,7	88,9	88,3
InceptionResNetV2	84,0	84,2	84,1
Squeezenet	86,4	85,4	85,9

Hasil ujicoba menggunakan arsitektur CNN dan optimasi gradient descent ditunjukkan pada Tabel 2,3 dan 4.

Tabel 2. Pengukuran performa klasifikasi citra retina dengan optimasi GDM

Arsitektur CNN	Sens. (%)	Spec. (%)	Acc. (%)	t
AlexNet	93,4	90,7	92,0	6'23
VGG16	90,7	94,0	92,31	35'5
VGG19	96,6	84,2	88,4	41'4
GoogLeNet	87,9	88,1	88,0	3'7
Inception-V3	81,5	85,4	83,3	2'26
ResNet50	86,8	88,2	87,5	2'4
ResNet101	94,7	82,3	87,5	7'30
IncResNetV2	94,2	75,1	82,1	28
Squeezenet	93,7	80,2	85,7	29'3

Ket : t (waktu) dalam satuan menit'detik

Hasil ujicoba menggunakan optimasi GDM menghasilkan nilai akurasi terbaik dengan arsitektur VGG16 yaitu 92,31%, sensitivitas 90,7% dan spesifisitas 94%.

Tabel 3. Pengukuran performa klasifikasi citra retina dengan optimasi RmsProp

Arsitektur CNN	Sens. (%)	Spec. (%)	Acc. (%)	t
AlexNet	94,7	83,0	88,0	6'21
VGG16	89,7	82,7	85,9	35'4
VGG19	85,0	82,7	83,8	42'2
GoogLeNet	77,9	90,9	83,2	1'15
Inception-V3	96,8	81,3	87,5	3'12
ResNet50	85,1	92,6	88,5	1'21
ResNet101	78,2	89,3	87,2	4'16
IncResNetV2	89,6	86,5	88,0	25
Squeezenet	94,7	83,0	88,0	18'3

Hasil ujicoba menggunakan optimasi RmsProp menghasilkan nilai akurasi terbaik dengan arsitektur ResNet50 yaitu 88,5%, sensitivitas 85,1% dan spesifisitas 92,6%.

Tabel 4. Pengukuran performa klasifikasi citra retina dengan optimasi Adam

Arsitektur CNN	Sens. (%)	Spec. (%)	Acc. (%)	t
AlexNet	93,5	88,3	90,7	6'43
VGG16	89,5	79,6	83,8	35'2
VGG19	92,6	79,6	84,9	44'2
GoogLeNet	97,0	84,2	89,6	54
Inception-V3	92,1	86,5	89,3	3'24
ResNet50	95,2	84,7	89,3	1'16
ResNet101	86,7	87,7	87,0	3'51
IncResNetV2	76,2	88,2	81,0	30
Squeezenet	83,2	89,5	86,1	12'3

Hasil ujicoba menggunakan optimasi Adam menghasilkan nilai akurasi terbaik dengan arsitektur AlexNet yaitu 90,7%, sensitivitas 93,5% dan spesifisitas 88,3%.

Pada ujicoba tahap 2 dilakukan inisialisasi learning rate 0,0001, max-epoch 6, validation frequency 3 dan minibatch-size 10. Nilai inisialisasi parameter tergolong rendah namun memiliki hasil akurasi lebih baik dibandingkan ujicoba tahap 1. Fungsi dari masing-masing parameter yaitu :

1. *Minibatch* akan mempengaruhi penggunaan memori selama pemrosesan. Ukuran minibatch yang lebih kecil membutuhkan lebih sedikit memori saat memproses.

2. *Max-epoch* adalah batasan jumlah iterasi yang dapat dilakukan untuk mengakses keseluruhan layer pada network.
3. *Validation Frequency* merupakan validasi dilakukan setelah menjalankan n data training. Misalnya ditentukan validation frequency n = 3, maka validasi dilakukan setelah menjalankan 3 iterasi data training.

Selanjutnya dibuat perbandingan antara hasil ujicoba pada tahapan pertama dan tahapan kedua. Perbandingan digunakan untuk menunjukkan bahwa optimasi berpengaruh terhadap peningkatan performa dari sistem klasifikasi. Hasil perbandingan akurasi ditunjukkan pada Tabel 5. Hasil perbandingan menunjukkan bahwa penggunaan algoritma optimasi tidak selalu berhasil pada setiap arsitektur CNN.

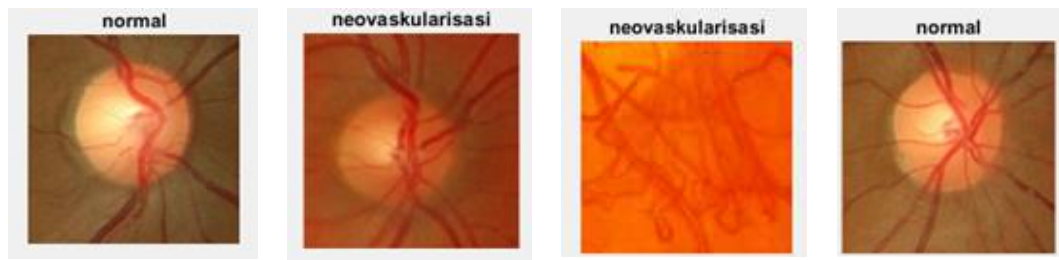
Pada arsitektur ResNet101, hasil akurasi tanpa optimasi menunjukkan prosentase lebih baik dibandingkan dengan hasil akurasi dengan metode optimasi. Namun, pada arsitektur CNN lain, akurasi lebih baik dapat dicapai dengan penggunaan algoritma optimasi *gradient descent*.

Tabel 5. Perbandingan hasil akurasi tanpa optimasi dan dengan optimasi

Arsitektur CNN	Akurasi Tahap 1 (%)	Akurasi GDM (%)	Akurasi RmsProp(%)	Akurasi Adam(%)
AlexNet	87,8	92,0	88,0	90,7
VGG16	88,8	92,3	85,9	83,8
VGG19	89,3	88,4	83,8	84,9
GoogLeNet	86,5	88,0	83,2	89,6
Inception-V3	83,5	83,3	87,5	89,3
ResNet50	88,0	87,5	88,5	89,3
ResNet101	88,3	87,5	87,2	87,0
InceptionResNetV2	84,1	82,1	88,0	81,1
Squeezenet	85,9	85,7	87,2	86,1

Hasil ujicoba pada tahap 2 menunjukkan akurasi terbaik adalah arsitektur VGG16 dengan 41 layer. VGG16 menggunakan konsep yang sama dengan VGG19. Perbedaan hanya terletak pada jumlah layer. VGG16 juga menggunakan konsep deeper layer. Semakin bertambah *layer* maka akurasi

semakin baik. Dimensi filter berukuran 3, 64, 128, 256, dan 512. Hasil klasifikasi ditampilkan pada Gambar 2. Gambar ini merupakan representasi hasil pengenalan dari klasifikasi fundus. Untuk perbandingan hasil dengan penelitian sebelumnya ditunjukkan pada Tabel 6.



Gambar 2. Representasi dari klasifikasi patch fundus. a, c dan d merupakan contoh klasifikasi benar sesuai label kategori. b. merupakan contoh klasifikasi salah

Tabel 6. Perbandingan hasil akurasi tanpa optimasi dan dengan optimasi

Author	Kelas	Jumlah Layer	Sens.(%)	Spec. (%)	Acc.(%)
Yu <i>et al.</i> [9]	2	16	88,85	96	91,92
Kwasigroch <i>et al.</i> [10]	2	16	-	-	81
Ghosh <i>et al.</i> [11]	2	29	-	-	85
<i>Proposed method</i>	2	47	90,4	88,2	89,3
		41	90,7	94	92,31

Ket : Sens. = Sensitivitas, Spec. = Spesifisitas, Acc = Akurasi

Perbandingan pada Tabel 6 menunjukkan capaian akurasi terbaik menggunakan VGG16 dengan 41 layer. Akurasi yang dicapai hingga 92,31% dengan Sensitivitas 90,7% dan Spesifisitas 94%.

SIMPULAN

Telah dilakukan klasifikasi citra fundus terhadap 2 kelas yaitu normal dan neovaskularisasi. Ujicoba telah dilakukan menggunakan 9 arsitektur CNN diantaranya AlexNet, VGG16, VGG19, GoogleNet, Inception-V3, ResNet50, ResNet101, InceptionResNet-V2, dan Squeezenet. Tahapan pertama dilakukan klasifikasi menggunakan arsitektur CNN dan SVM. Hasil ujicoba menunjukkan akurasi hingga 89,3% dengan sensitivitas 90,4% dan spesifisitas 88,2% menggunakan arsitektur VGG19.

Pada tahapan kedua dilakukan klasifikasi citra fundus menggunakan sembilan arsitektur CNN dengan teknik transfer learning dan metode optimasi gradient descent. Hasil ujicoba menunjukkan akurasi hingga 92,31% dengan sensitivitas 90,7% dan spesifisitas 94% menggunakan arsitektur VGG16. Hal ini menunjukkan bahwa ujicoba menggunakan algoritma optimasi

menghasilkan akurasi lebih baik dibandingkan tanpa optimasi.

SARAN

Penelitian selanjutnya dapat dikembangkan arsitektur CNN berdasarkan arsitektur VGG dengan jumlah layer dan parameter lebih sederhana namun lebih baik dalam hal akurasi hasil klasifikasi.

DAFTAR PUSTAKA

- [1] Lechun, "gradient based learning applied to document recognition," *proc.of IEEE*, no. november, pp. 1–46, 1998.
- [2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," 2012.
- [3] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in *ICLR*, 2015, pp. 1–14.
- [4] C. Szegedy, V. Vanhouke, S. Ioffe, and J. Shlens, "Rethinking the Inception Architecture for Computer Vision Christian," *IEEE Explor.*, pp. 2818–2826,

- 2015.
- [5] K. He and J. Sun, "Deep Residual Learning for Image Recognition," *IEEE Xplore*, pp. 1–9, 2015.
- [6] C. Szegedy, S. Ioffe, and V. Vanhoucke, "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning," pp. 1–12, 2016.
- [7] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size," in *ICLR*, 2017, pp. 1–13.
- [8] W. Setiawan, M. . Utoyo, and R. Rulaningtyas, "Classification of neovascularization using convolutional neural network model," *TELKOMNIKA*, vol. 17, no. 1, pp. 463–473, 2019.
- [9] S. Yu, D. Xiao, and Y. Kanagasingam, "Exudate Detection for Diabetic Retinopathy With Convolutional Neural Networks," pp. 1744–1747, 2017.
- [10] A. Kwasigroch, B. Jarzembinski, and M. Grochowski, "Deep CNN based decision support system for detection and assessing the stage of diabetic retinopathy," *2018 Int. Interdiscip. PhD Work.*, pp. 111–116, 2018.
- [11] R. Ghosh, "Automatic Detection and Classification of Diabetic Retinopathy stages using CNN," 2017.
- [12] E. Decenci re *et al.*, "Feedback On A Publicly Distributed Image Database : The Messidor Database," *Image Anal Stereol*, vol. 33, pp. 231–234, 2014.
- [13] S. Ruder, "An overview of gradient descent optimization," 2017.
- [14] G. E. Hinton, N. Srivastava, and K. Swersky, "Lecture 6a-overview of mini-batch gradient descent," *COURSERA Neural Networks Mach. Learn.*, p. 31, 2012.
- [15] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in *ICLR*, 2015, pp. 1–15.